Heart Disease Prediction Using Machine Learning

Prof. Anuja Phapale
Department of Information Technology
All India Shri Shivaji Memorial
Society's Institute of Information Technology
Pune, India

Arshiya Mulla
Department of Information Technology
All India Shri Shivaji Memorial
Society's Institute of Information Technology
Pune, India
arsiyamull@gmail.com

Kaveri Kumavat
Department of Information Technology
All India Shri Shivaji Memorial
Society's Institute of Information Technology
Pune, India
kaverikumava@gmail.com

Smita Ingavale
Department of Information Technology
All India Shri Shivaji Memorial
Society's Institute of Information Technology
Pune, India
smitaaingawale109@gmail.com

Apeksha Jadhav
Department of Information Technology
All India Shri Shivaji Memorial
Society's Institute of Information Technology
Pune, India
apekshajadh@gmail.com

ABSTRACT

In today's modern world cardiovascular disease is the most lethal one. This disease attacks a person so instantly that it hardly gets any time to get treated with. So diagnosing patients correctly on timely basis is the most challenging task for the medical fraternity. A wrong diagnosis by the hospital leads to earn a bad name and loosing reputation. At the same time treatment of the said disease is quite high and not affordable by most of the patients particularly in India. The purpose of this paper is to develop a cost effective treatment using data mining technologies for facilitating data base decision support system. Almost all the hospitals use some hospital management system to manage healthcare in patients. Unfortunately most of the systems rarely use the huge clinical data where vital information is hidden. As these systems create huge amount of data in varied forms but this data is seldom visited and remain untapped. So, in this direction lots of efforts are required to make intelligent decisions. The diagnosis of this disease using different features or symptoms is a complex activity. In this paper using varied data mining technologies an attempt is made to assist in the diagnosis of the disease in question.

Keywords: cardiovascular disease, data mining, intelligent decisions, symptoms

1. INTRODUCTION

Today, many hospitals manage healthcare data using healthcare information system; as the system contains huge amount of data, used to extract hidden information for making intelligent medical diagnosis. The value of machine learning in healthcare is its ability to process huge datasets beyond the scope of human capability, and then reliably convert analysis of that data into clinical insights that aid physicians in planning and providing care, ultimately leading to better outcomes, lower costs of care, The main objective of this research is to build Intelligent Heart Disease Prediction System that gives diagnosis of heart disease using historical heart database. To develop this system, medical terms such as sex, blood pressure, and cholesterol like 13 input attributes are used. To get more appropriate results, two more attributes i.e. obesity and smoking are used, as these attributes are considered as important attributes for heart disease. The data mining classification techniques viz. Neural Networks, Decision Trees.Random Forest, and Naive Bayes are used.

2. LITERATURE SURVEY

Very few systems use the available clinical data for prediction purposes and even if they do, they are restricted by the large number of association rules that apply. Diagnosis of the condition solely depends upon the Doctors' intuition and patient's records. Detection is not possible at an earlier stage.

In the existing system, practical use of various collected data is time consuming. There are only few decision support systems available in medical industry whose functionalities are very limited.

Table 1. Literature Survey

SR.NO	TITLE	PUBLICATION YEAR	AUTHOR	ALGORITHM	KEYWORDS
1.	"Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review	**2016**	V. Krishnaiah, G. Narasimha, N. Subhash Chandra,	It focus on various data mining practices that are valuable in heart disease forecast.	WHO, distress,Organization,data mining
2.	"Study of Heart Disease Prediction using Data Mining"	**2016**	K.Sudhakar, Dr. M. Manimekalai	The major challenge that the Healthcare industry faces now-a-days is superiority of facility	Healthcare, Diagnosing
3.	An Improved Method for Disease Prediction using Fuzzy Approach	**2015**	NagannaChetty, Kunwar Singh Vaisla, NagammaPatil	It discuss symptoms of heart disease which may vary accordingly	back pain,jaw pain, neck pain, stomach disorders,and tininess of breath,chest pain, arms and shoulders pains.
4.	Prediction of Heart Disease Using Machine Learning	**2019**	Aditi Gavhane; Gouthami Kokkula; Isha Pandya; Kailas Devadkar	Predicts the risk of heart disease with an accuracy of 89%. Time for optimization high, very slow Convergence Algo: Random Forest	Conferences, Prediction algorithms, Machine learning, Sensors

5.	Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques	**2019**	Senthilkumar Mohan; Chandrasegar Thirumalai; Gautam Srivastava	Accuracy of heart disease dataset improved 5% over classification algorithm. Algo:KNN	Diseases,Heart, Data mining, Support vector machines,Feature extraction, Predictive models
6.	Human heart disease prediction system using data mining techniques	**2016**	J. Thomas; R Theresa Princy	AKNN takes more time to train data, Does not work if data is noisy.	band-pass filters, biometrics (access control),

3. METHDOLOGY

The Heart Disease Prediction application is an end user support and consultation project. Here, we propose a application that allows users to get instant guidance on their heart disease through an intelligent system. The application is fed with various details and the heart disease associated with those details. The application allows user to share their heart related issues. It then processes user specific details to check for various illness that could be associated with it. Here we use some intelligent data mining techniques to guess the most accurate illness that could be associated with patient's details. Based on result, the can contact doctor accordingly for further treatment. In general, the more trees in the forest the more robust the forest looks like. In the same way in the random forest classifier, the higher the number of trees in the forest gives the high accuracy results.

Module 1 User module:

In this module patient and doctor will register themselves.Once registration is done,patient can take appointment ,post query and chat with the doctor.Once doctor is register he can post comment and can chat with patient.

Module 2 Report module:

In this module Patient report will be generated in doc format.

Module 3 Clustering module:

- Randomly select "k" features from total "m" features.
  - Where k << m
- Among the "k" features, calculate the node "d" using the best split point.
- Split the node into daughter nodes using the best split.
- Repeat 1 to 3 steps until "l" number of nodes has been reached.
- Build forest by repeating steps 1 to 4 for "n" number times to create "n" number of trees.

4. DESIGN

4.1 Hardware

- Computer or Laptop
- 4GB RAM

4.2 Software

- OpenCV
- Python
- TensorFlow
- SMTP Email

5. SYSTEM ARCHITECTURE

The general diseases prediction system predicts chance of presence of a diseases present in a patient on the basis of their symptoms. It will also recommend necessary precautionary measures required to treat the predicted diseases. The system will initially be fed data from different sources i.e. patients, the data will then be pre-processed before further process is carried out, this is done so as to get clean data from the raw initial data, as the raw data would be noisy, or flawed. This data will be processed using Data minings algorithms, the system, will be trained so as to predict the diseases based on the input data given by the user.

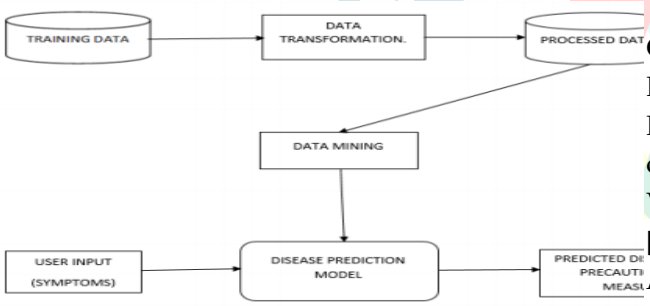

Fig 1.System Architecture

Advantage:

1. User can search for doctor's help at any point of time.
2. User can talk about their Heart Diseases and get instant diagnosis.
3. Doctors get more clients.
4. Very useful in case of emergency.

6. CONCLUSION

The overall objective of our work is to predict more accurately the presence of heart disease. In this topic two more input attributes obesity and smoking are used to get more accurate results. Data mining classification techniques were applied namely Random Forest. This technique will help us to achieve 97% accuracy as per algorithm

REFERENCES

[1] J Thomas MR, Lip GY. Novel risk markers and risk assessments for cardiovascular disease. Circulation research. 2017; 120(1):133–149. https://doi.org/10.1161/CIRCRESAHA.116.309955 PMID: 28057790

[2] Ahmed M. AlaaID1, Thomas Bolton, Emanuele Di Angelantonio, James H.F. RuddID, Mihaela van der Schaar,―Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants‖, PLOS ONE 14(5): e0213653. https://doi.org/10.1371/journal, May 15, 2019H. Poor, ―A Hypertext History of Multiuser Dimensions,‖ MUD History, http://www.ccs.neu.edu/home/pb/mudhistory.html. 1986. (URL link *include year)

[3] Stephen F. Weng, Jenna Reps, Joe Kai1, Jonathan M. Garibaldi, Nadeem Qureshi, ―Can machine-learning improve cardiovascular risk prediction using routine clinical data?‖, PLOS ONE | https://doi.org/10.1371/journal.pone. 0174944 April 4, 2017

[4] Rine Nakanishi, Damini Dey, Frederic Commandeur, Piotr Slomka, ―Machine Learning in Predicting Coronary Heart Disease and Cardiovascular Disease Events: Results from The Multi-Ethnic Study of Atherosclerosis (Mesa)‖, JACC Mar- 20, 2018, Volume 71, Issue 11

[5] https://www.cdc.gov/heartdisease/facts.htm. Available [Online].

[6] Senthilkumar Mohan, Chandrasegar Thirumalai, Gautam Srivastava ―Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques‖, Digital Object Identifier 10.1109/ACCESS.2019.2923707, IEEE Access, VOLUME 7, 2019 S.P. Bingulac, ―On the Compatibility of Adaptive Controllers,‖ Proc. Fourth Ann. Allerton Conf. Circuits and Systems Theory, pp. 8-16, 1994. (Conference proceedings)